TEXT GENERATION AND SEARCHING METHOD AND SYSTEM

RELATED APPLICATIONS

[0001] The following U.S. patent is hereby incorporated by reference into the subject application as if set forth herein in full: (1) U.S. Patent No. 6,463,417, entitled "Method of Distributing Health Information".

FIELD OF THE INVENTION

[0002] This invention relates to database record searching, and, more particularly, to text-based searching of database records.

BACKGROUND

[0003] The efficient management of large sets of computer-based data is a difficult task. In addition to the physical hardware requirements needed to effectuate the storage of the data, once the data is stored, the management and organization of the data may prove daunting.

[0004] Databases are often used to manage and maintain large sets of data, such that the data is organized around a defined database structure. When retrieving data stored within the database, the individual records of the database must be searched. Unfortunately, as the number of records within the database increases, the search time associated with retrieving the data increases dramatically, which often results in unacceptable delay times and latency.

SUMMARY OF THE INVENTION

[0005] According to a first implementation, a text-generation method includes receiving data records, such that each data record includes one or more data fields and a field value associated with each data field. A text-string is generated for each data record,

1

such that each text-string includes one or more text-based data descriptors. Each data descriptor includes a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field.

[0006] One or more of the following features may also be included. The text-strings may be stored as a text-based data file, such as an ASCII file. Each text-string may include a record identifier that identifies the data record to which the text-string is related. Each data descriptor may include one or more starting characters, one or more separator characters, and one or more ending characters. The field descriptor may be positioned between the separator characters and one of the starting characters and the ending characters. The value descriptor may be positioned between the separator characters and the other of the starting characters and the ending characters. The data records may be representative of the medical records of patients.

[0007] According to a further implementation, a search method includes defining a first target value for each of one or more data fields within a database record structure of a database. The database includes a plurality of data records. A plurality of text-strings are searched, such that each text string is associated with one of the data records and includes one or more text-based data descriptors.

[0008] Each data descriptor includes a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field.

[0009] A first result set is generated by identifying one or more text-strings that include a value descriptor that is essentially equivalent to at least one of the first target values.

[0010] One or more of the following features may also be included. The first target values may include one or more wildcard descriptors. The data record associated with one or more of the text-strings identified in the first result set may be retrieved.

[0011] A second target value may be defined for each of one or more data fields within

the database record structure of the database. The plurality of data records included in the database may be searched, and a second result set may be generated by identifying one or more data records that include a field value that is essentially equivalent to at least one of the second target values. One or more of the data records identified in the second result set may be retrieved.

[0012] According to a further implementation, a computer program product resides on a computer readable medium on which a plurality of instructions are stored. When executed by the processor, the instructions cause that processor to: receive data records, such that each data record includes one or more data fields and a field value associated with each data field; and generate a text-string for each data record, such that each text-string includes one or more text-based data descriptors. Each data descriptor includes: a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field.

[0013] According to a further implementation, a computer program product resides on a computer readable medium on which a plurality of instructions are stored. When executed by the processor, the instructions cause that processor to: define a first target value for each of one or more data fields within a database record structure of a database, such that the database includes a plurality of data records; search a plurality of text-strings, wherein each text string is associated with one of the data records and includes one or more text-based data descriptors. Each data descriptor includes: a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field. A first result set is generated by identifying one or more text-strings that include a value descriptor that is essentially equivalent to at least one of the first target values.

[0014] According to a further implementation, a searching system includes a server system having a computer processor and associated memory, the server system including a

3

database that includes a plurality of data records. The server system is configured to: define a first target value for each of one or more data fields within a database record structure of the database and search a plurality of text-strings. Each text string is associated with one of the data records and includes one or more text-based data descriptors, such that each data descriptor includes: a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field. A first result set is generated by identifying one or more text-strings that include a value descriptor that is essentially equivalent to at least one of the first target values.

[0015] According to a further implementation, a data structure includes a database having a plurality of data records, such that each data record includes one or more data fields, and a field value is associated with each data field. The data structure includes a text-string for one or more data records, such that each text-string includes one or more text-based data descriptors. Each data descriptor includes: a field descriptor that defines a specific data field within the data record to which the text-string is related, and a value descriptor that defines the field value associated with the specific data field.

[0016] The details of one or more implementations is set forth in the accompanying drawings and the description below. Other features and advantages will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagrammatic view of a record organization system coupled to a distributed computing network;

FIG. 2 is a more-detailed diagrammatic view of the record organization system of FIG. 1;

FIG. 3 is a diagrammatic view of a key maintenance module and a key processing module of the record organization system of FIG. 1;

4

FIG. 4 is a diagrammatic view of a record processing module of the record organization system of FIG. 1;

FIG. 5 is a diagrammatic view of a medical record;

FIG. 6 is a diagrammatic view of a patient selection display screen rendered by the record organization system of FIG. 1;

FIG. 7 is a diagrammatic view of a record searching module of the record organization system of FIG. 1; and

FIG. 8 is a diagrammatic view of a query definition display screen rendered by the record organization system of FIG. 1.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0017] Referring to FIG. 1, there is shown a record organization system 10 that manages the various access keys 12, 14, 16 possessed by a medical service provider 18. Access keys 12, 14, 16 allow the medical service provider 18 to access the medical records (not shown) of various patients 20, 22, 24 (respectively).

[0018] Record organization system 10 typically resides on and is executed by a computer 26 that is connected to a network 28. Computer 26 may be a web server running a network operating system, such as Microsoft Window 2000 Server [tm], Novell Netware [tm], or Redhat Linux [tm]. Typically, computer 26 also executes a web server application, such as Microsoft IIS [tm], Novell Webserver [tm], or Apache Webserver [tm], that allows for HTTP (i.e., HyperText Transfer Protocol) access to computer 26 via network 28.

[0019] The instruction sets and subroutines of record organization system 10, which are typically stored on a storage device 30 coupled to computer 26, are executed by one or more processors (not shown) and one or more memory architectures (not shown) incorporated into computer 26. Storage device 30 may be, for example, a hard disk drive, a tape drive, an optical drive, a RAID array, a random access memory (RAM), or a read-only memory (ROM).

[0020] As will be explained below in greater detail, a patient (e.g., patient 20) typically provides a key (e.g., access key 12) to medical service provider 18 through a patient computer 32, which is also connected to network 28. Additionally, medical service provider 18 accesses record organization system 10 through a client computer 34.

[0021] Referring also to FIG. 2, record organization system 10 includes a centralized key repository 50 and a centralized medical records repository 52. Additionally, record organization system 10 includes a key maintenance module 54, a key processing module 56, a record processing module 58, and a record searching module 60, each of which will be discussed below in greater detail.

[0022] Centralized medical records repository 52 allows for the centralized storage of medical records 62, 64, 66 concerning various patients 20, 22, 24 respectively. As disclosed in U.S. Patent No. 6,463,417, medical records 62, 64, 66 are typically divided into portions or levels, in that certain portions are considered more confidential than other portions. For example, a portion / level of the medical record that may be considered the least confidential might include general patient identification information and information concerning the patient's blood type and allergies. A portion / level of a medical record that may be considered to have an intermediate level of confidentiality might include information concerning the serological data, psychiatric data, cardiology data, and genetic data. A portion / level of the medical record that may be considered highly confidential may include infectious disease (e.g., HIV, and sexually transmitted diseases) data.

[0023] This specific assignment of confidentiality levels and the apportionment of the medical record into various portions / levels is for illustrative purposes only and is not intended to limit the scope of this disclosure.

[0024] Medical records 62, 64, 66 may be incrementally generated / configured online by the various medical service providers that provide care to patients 20, 22, 24. Alternatively, existing medical records may be uploaded (i.e., transferred) to medical records repository 52 from a remote storage location (not shown).

[0025] Referring also to FIG. 3, patients 20, 22, 24 use key maintenance module 54 to generate 100 access keys 12, 14, 16 that grant access to various portions of the respective medical records 62, 64, 66. Accordingly, though the use of key maintenance module 54, the patient can generate access keys that not only regulate who has access to their medical records, but also regulate the level of access (i.e., which portions of a patient's medical record are viewable by the medical service provider to which the key is provided). Examples of access keys 12, 14, 16 are passwords (that allow access to various portions of a medical record) and decryption keys (that decrypt various portions of an encrypted medical record).

[0026] Typically, key maintenance module 54 is a web-enabled application that is accessed by the patients (e.g., patient 20) through a browser application (e.g., Microsoft Internet Explorer ^{tm}, or Netscape Navigator ^{tm}) that is running on patient computer 32. Alternatively, key maintenance module 54 may be a local application that is executed locally on patient computer 32.

[0027] As stated above, key maintenance module 54 allows a patient to generate 100 an access key for a specific medical service provider that grants, to that medical service provider, a defined level of access to that patient's medical records. Once this access key is generated, the access key is transmitted 102 to the medical service provider 18. This transmission of the access key may be implemented by transferring the access key from the patient to the medical service provider. This may occur by attaching the access key to an email that is transmitted to the medical service provider. Once received, the medical service provider may then transfer the newly-generated key to the key processing module 56 (to be discussed below in greater detail) of the record organization system 10. Alternatively, the patient may directly transfer the newly-generated key to the key processing module 54 of the record organization system 10.

[0028] Regardless of the manner in which the patient transfers the access key to the medical service provider, the access key will ultimately be received 120 by key processing

7

module 56, which receives any access keys (e.g., keys 12, 14, 16) generated and transmitted by patients 20, 22, 24. Once these keys are received 120, they are stored 122 on centralized key repository 50. Additionally, if record organization system 10 is servicing multiple medical service providers (e.g., medical service providers 17, 18, 19), the received keys are associated 124 with the appropriate medical service provider so that the keys transmitted to a first provider are not available to a second provider.

[0029] Referring also to FIG. 4, when medical records (i.e., uploaded existing records, newly-generated records, and/or amended records) are initially received 138, record processing module 58 stores 140 the medical record on centralized medical record repository 52. Typically, medical record repository 52 is a database that allows for the organized storage and retrieval of the medical records 62, 64, 66.

[0030] Once these medical records are stored on medical record repository 52, record processing module 58 allows the medical service provider 18 to access 142 the medical records 62, 64, 66 stored on medical records repository 52. However, the medical service provider 18 is only given access to the portions of the medical records for which the medical service provider 18 possesses the appropriate key. For example, assume that medical service provider 18 is a medical clinic that provides an array of medical services to its patients. Further, assume that patient 20 uses medical service provider 18 for all of their medical needs; patient 22 uses medical service provider 18 solely for treatment of depression; and patient 24 uses medical service provider 18 solely for treatment of HIV.

[0031] Concerning the access keys generated by each of these patients for medical service provider 18: patient 20 would typically provide medical service provider 18 with an access key (i.e., key 12) that grants access to their entire medical record; patient 22 would typically provide medical service provider 18 with an access key (i.e., key 14) that grants access to the general and psychiatric portions of their medical record; and patient 22 would typically provide medical service provider 18 with an access key (i.e., key 16) that grants access to the general and infectious disease portions of their medical record.

8

[0032] Record processing module 58 is typically a web-enabled application that is accessed by the medical service provider 18 through a browser application (e.g., Microsoft Internet Explorer [tm], or Netscape Navigator [tm]) that is running on client computer 34. Typically, medical service provider 18 logs into key organization system 10 using an encrypted SSL (i.e., secure sockets layer) connection.

[0033] Referring also to FIG. 5, medical records 62, 64, 66 are typically database records 160 that contain data fields (e.g., data field 162), each of which includes a field name 164 and a field value 166. Additionally, as discussed above, medical records 62, 64, 66 includes serological data 168, psychiatric data 170, cardiology data 172, genetic data 174, and infectious disease data 176, each of which may be further broken down into data fields.

[0034] To enhance the searchability of centralized medical record repository 52, record processing module 58 may process each record to generate 144 a text string that relates to that record. A example of a text string for record 160 is:

> **<first_name:timothy> <last_name:smith> <ss#:123456789> <street:20 elm street> <city:boston> <state:ma> <zip:02110> <gender:male> <weight:183> <cholosterol:172> <sys_press:12)> <dia_press:70> <data_record: 1243562>**

[0035] The above-listed text string is a textual representation of the field-based data included within the "patient identification portion" of the medical record 62 of patient 22 (i.e., Timothy Smith).

[0036] The text string includes one or more data descriptors (e.g., **<dia_press:70>**), each of which includes a field descriptor (e.g., **dia_press**) and a field value (e.g., **70**). The field descriptor and the field value are typically separated by a separator character (e.g., :), and the data descriptor begins with beginning characters (e.g., <) and ends with ending characters (e.g., >).

[0037] The field descriptor defines a specific data field within the data record to which the text-string is related. For example, the field descriptor **"dia_press"** relates to data field

9

164 in record 62. Further, the value descriptor defines the field value associated with the same specific data field. For example, the value descriptor "**70**" relates to field value 166

[0038] The text string typically also includes a record identifier (e.g., **<data_record: 1243562>**) that associates the text-string with the medical record upon which the text-string is based, so that the medical record may be subsequently retrieved.

[0039] The text-strings generated by record processing module 58 are typically stored 146 within centralized medical record repository 52. The text-strings may be stored in a group as a single text file, such as an ASCII (i.e., American Standard Code for Information Interchange) file. Alternatively, the text-strings may be stored as individual text-based files. Regardless of the manner in which the text-strings are stored, as will be discussed below, by searching the text strings for the occurrence of specific data descriptors, the medical records that contain desired information may be quickly identified.

[0040] Referring also to FIG. 6, when accessing record organization system 10, record processing module 58 provides the medical service provider 18 with a rendered screen display 180 that includes a list of patient identifiers 182. Patient identifiers 182 define the particular patient(s) who provided access keys to medical service provider 18 (i.e., granting medical service provider 18 access to various portions of their medical record(s)). The patient identifiers 182 may be any element that uniquely identifies the patient, such as the patient's name, the patient's social security number, or a unique patient number. In this particular example, Mary Jones is patient 20, Timothy Smith is patient 22, and James Greco is patient 24.

[0041] The presence of each of these names in the list of patient identifiers 182 indicates that a key was received from that patient. In order to access the medical record of a patient for which the medical service provider has a key (i.e., for one of the patients listed in the list of patient identifiers 182), the medical service provider 18 selects the appropriate identifier using a mouse pointer 184 (or some other pointing device, not shown). For example, if the medical service provider wanted to access the medical record of Timothy

10

Smith (i.e., patient 22), medical service provider 18 would typically click (using a mouse) on the specific identifier 186 associated with Timothy Smith. Record processing module 58 would then, in turn, use access key 14 to access (i.e., retrieve, decrypt, and display) medical record 64, the medical record of Timothy Smith, i.e., patient 22.

[0042] Medical record 64 may be displayed in a separate window or displayed full screen on the display of client computer 34. As discussed above, the key provided to the medical service provider 18 only allows access to the portion(s) of the patient's medical record that the patient wishes to allow access. As discussed above, Timothy Smith (i.e., patient 22) is being treated by medical service provider 18 for depression and access key 14 grants access to the general and psychiatric portions of Timothy Smith's medical record, such that a link (e.g., link 188) to each available portion is displayed on the right-hand side of medical record 62. However, access key 14 does not permit access (i.e., prohibits access) to the other portions of Timothy Smith's medical record, namely Allergies, Serological Data, Cardiology Data, Genetic Data, and Infectious Disease Data. Accordingly, the links (e.g., link 190) to the unavailable data portions are struck-through. Other methods of differentiating the available portions from the unavailable portions of a medical record may be used, such as graying-out or not displaying links to the unavailable portions.

[0043] By clicking on the links to the available portions of the medical record, a specific available portion is displayed by record processing module 58.

[0044] Referring also to FIG. 7, record searching module 60 facilitates the searching of medical records stored on the centralized medical record repository 52. This searching may be performed by a medical service provider (e.g., medical service provider 18) or an administrator of record organization system 10. However, the party doing the searching may only search the portions of the medical records to which they are granted access. This access may be based on the possession of keys (i.e., for medical service providers) or administrative privileges (i.e., for administrators), for example.

[0045] Record searching module 60 allows a user (e.g., a medical service provider, an administrator, etc.) to define 200 a target value for one or more of the data fields (e.g., data field 162) within the database record structure of the medical records (e.g., medical record 62) stored on centralized medical record repository 52. The database record structure refers to the field structure of a database records. For example, database record 62 includes twelve specific data fields (namely, first name, last name, social security number, street, city, state, zip code, gender, weight, cholesterol, systolic pressure, and diastolic pressure) and five data portions (namely serological data, psychiatric data, cardiology data, genetic data, and infectious disease data). As stated above, the patient controls the access to the various portions of their medical record through the use of access keys. Additionally, these portions are typically subdivided into numerous data fields.

[0046] Typically, record searching module 60 is a web-enabled application that is accessed by the user (e.g., medical service provider 18) through a browser application (e.g., Microsoft Internet Explorer ^{tm}, or Netscape Navigator ^{tm}) that is running on a computer (e.g., client computer 34). Alternatively, record searching module 60 may be a locally-executed application.

[0047] Referring also to FIG. 8, when a user initiates a search, a query definition display 240 is rendered by record searching module 60 that allows the user to define 200 the search terms (i.e, target values) of a database record query. For example, if medical service provider 18 wished to search the medical records (to which they have access) to determine which of their patients were male, medical service provider 18 would enter the search term "male" into field 242 of display 240. These search terms may be whole words, portions of words, and/or wildcard descriptors (e.g., *).

[0048] Medical service provider 18 would then initiate 202 the search by "clicking" on the text-string search button 244 with mouse pointer 246. As, in the example, medical service provider 18 only has access to the medical records of three patients, namely Mary Jones (i.e., patient 20), Timothy Smith (i.e., patient 22), and James Greco (i.e., patient 24),

12

the result list (not shown) for the query would include two entries, namely Timothy Smith and James Greco.

[0049] When executing a text-string based search, the search is executed by associating the search term(s) entered (i.e., "male") with the data field into which that search term was entered (i.e., "gender). This, in turn, generates a search term in the form of a data descriptor (i.e., <gender:male>). Therefore, when executing the text-string search, the text string for each medical record available to the initiator of the search is examined to determine if any of these text strings contain the search term data descriptor. As discussed above, the text-string associated with patients Timothy Smith (i.e., patient 22) and James Greco (i.e., patient 24) include the data descriptor <gender:male>. Additionally, each text-string includes a record identifier that associates the text-string to the medical record on which it is based. Therefore, once the relevant text strings are determined, the related medical records are easily ascertained. By performing text-based searches on text-strings, the search speed is increased and latency is reduced.

[0050] The result list (not shown) generated 204 in response to this query includes the names of Timothy Smith and James Greco. Typically, the result list is in the form of an HTML document. As the patient names are typically embedded links, when the user clicks on one or more of these links, the corresponding medical record is retrieved 206.

[0051] In addition to the text-based searches performed above, a record-based search may also be performed. Typically, a text-based search is performed as an initial search to generate a first result set. Often, this result set is quite large and, therefore, a secondary search may be performed on the first result set. Accordingly, once a determination 208 is made that the first result set needs to be further reduced in size, a secondary record-based search may be performed by defining 210 additional search terms (i.e., target values) that further restrict the result set. The new search is then performed 212 and a refined result set is generated 214 and presented to the user.

[0052] Assuming that James Greco weighs 230 pounds, if the additional search term defines the desired weight as "<200 pounds", Timothy Smith will be the only patient specified in the second result set, as he weighs 183 pounds. As above, the medical record(s) specified in the second result set are easily retrieved 216 by clicking on the appropriate link. The search terms may be repeatedly refined until the result set is reduced to a size acceptable to the user.

[0053] As stated above, the secondary query may be performed using traditional record-based searching technique employed by database search engines. If record-based searching is desired, a record based search is initiated 212 by "clicking" on the record-based search button 248 with mouse pointer 246.

[0054] While medical record 66 is shown to include a plurality of links to the available portions of the medical record, other configurations are possible. For example, when clicking on a specific identifier (e.g., identifier 164), a medical record may be displayed that only includes the portions to which the medical service provider has access.

[0055] While the centralized key repository 50 and the centralized medical record repository 52 are described above as being located on a remote server, other configurations are possible. For example, as is known in the art, one or more of these repositories may be distributed across multiple computers / servers.

[0056] A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made. Accordingly, other implementations are within the scope of the following claims.

14